Assessment of height accuracy of the DEM for species habitat analysis and modeling

Haller, R M ; Imfeld, S

Abstract: In mountainous areas, the use of DEM as environmental parameter in local and global models is widely accepted. Derived information like slope, aspect and others are often used as substitutes for missing data on specific habitat parameters to model the probability of species occurrence or absence. For habitat analysis of species, point observations were combined with underlying terrain information. The use of DEM with a resolution of 25 m is still accepted as sufficient. Different studies have shown the influence of specific topographic elements like ridges, depression, steep and flat areas on species distributions. Data providers often provide an overall accuracy description for the total area. The spatial correlation of error is hardly ever described, although this influences notably the derived data from DEM and could consequently be a major source of uncertainty in models using such derivatives. We investigated the error of neighbouring cells of two different DEMs with a resolution of 20 m and 25 m, derived from photogrammetry and digitized from contour lines, respectively. 5 sampling areas each with 5*5 control points of the precise lattice coordinate of the DEM were established in the area of the Swiss National Park to investigate error correlations in open and forested as well as in planar and steep areas. Moreover, 246 survey points of a slightly irregular grid in the open area were used. The reference points were measured with surveying techniques. The results show a spatial correlation of the error on all plots as well as significant differences between the different subareas. Nevertheless, the average difference of error between neighbouring points is 1.01 m. The derived mean slope error from cell to cell (20m) is therefore 2.9°. The maximum error between two neighbouring cells of 25 m is 8.14m which is resulting in a slope mismeasurement of 19.0°. These results show the importance of the knowledge about spatial correlation of error in DEM, especially its consequences on derived data of DEM used in more complex model calculations such as species habitat models.

Originally published at:
Haller, R M; Imfeld, S (2007). Assessment of height accuracy of the DEM for species habitat analysis and modeling. In: 5th International Symposium Spatial Data Quality 2007. Theme: Modelling qualities in space and time, Enschede (NL), 13 June 2007 - 15 June 2007. ISPRS, online.

ASSESSMENT OF HEIGHT ACCURACY OF DEM FOR SPECIES HABITAT ANALYSIS AND MODELLING.

R. M. Haller [a, *], S. Imfeld [b]

[a] Swiss National Park, Chasa dal Parc, CH-7550 Zernez, Switzerland - rhaller@nationalpark.ch
[b] Department of Geography, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland – imfeld@geo.unizh.ch

KEY WORDS: DEM accuracy, spatial correlation of error, slope, Swiss National Park, GIS

ABSTRACT:

In mountainous areas, the use of DEM as environmental parameter in local and global models is widely accepted. Derived information like slope, aspect and others are often used as substitutes for missing data on specific habitat parameters to model the probability of species occurrence or absence. For habitat analysis of species, point observations were combined with underlying terrain information. The use of DEM with a resolution of 25 m is still accepted as sufficient. Different studies have shown the influence of specific topographic elements like ridges, depression, steep and flat areas on species distributions. Data providers often provide an overall accuracy description for the total area. The spatial correlation of error is hardly ever described, although this influences notably the derived data from DEM and could consequently be a major source of uncertainty in models using such derivatives.

We investigated the error of neighbouring cells of two different DEMs with a resolution of 20 m and 25 m, derived from photogrammetry and digitized from contour lines, respectively. 5 sampling areas each with 5*5 control points of the precise lattice coordinate of the DEM were established in the area of the Swiss National Park to investigate error correlations in open and forested as well as in planar and steep areas. Moreover, 246 survey points of a slightly irregular grid in the open area were used. The reference points were measured with surveying techniques.

The results show a spatial correlation of the error on all plots as well as significant differences between the different subareas. Nevertheless, the average difference of error between neighbouring points is 1.01 m. The derived mean slope error from cell to cell (20m) is therefore 2.9°. The maximum error between two neighbouring cells of 25 m is 8.14m which is resulting in a slope mismeasurement of 19.0°.

These results show the importance of the knowledge about spatial correlation of error in DEM, especially its consequences on derived data of DEM used in more complex model calculations such as species habitat models.

1. INTRODUCTION

Habitat models at global and regional scales became popular in recent years. The use of GIS redefined the assessment of the needs of plants and animal species for conservation and management issues (Selkirk 2002). Different approaches like habitat suitability indices (HSI) or niche analysis have been adopted in GIS environments (Mladenoff and Sickley 1997, Filli et al. 2000, Hirzel et al. 2002, Li et al. 2002). Animal behaviour has been analysed and explained by incorporating spatial data bases of environmental information as inputs for statistical methods.

For an exhaustive and meaningful conclusion for conservation purposes, the set of possible variables is chosen as numerous as possible (Corsi et al. 1999, McLoughlin et al. 2002, Campell 2003). By contrast, the available spatial data are often limited in terms of availability and accuracy and the specific parameters are often missing. Therefore, the missing variable originally intended for use is derived from secondary data (Guisan and Zimmermann 2000). For example, the variable “forest” is often used as remote area for animals with food resources and privacy. In mountainous regions, the DEM are widely used for different variables like climatic conditions (height, shadow, exposition), protection (as steeper as less accessible for enemies), to define upper and lower limits for food resources, to define good or bad thermal conditions for birds of prey or vultures or more general as lack of disturbance due to the missing human influence. Some studies did use the direct factor “slope” or “exposition” in their studies to explain animal occurrence or absence.

The quality of the spatial data in general and the DEM in particular are almost never checked and approved for the fitness of use. At best the users rely on the indications given by data owner and producer. Additionally, it is widely accepted that it is the manufacturers purpose to describe the quality of DEM sufficiently (Aguilar et al. 2007).

For DEMs numerous studies have been presented describing various aspects of the quality (Maune 2001, Shi et al. 2002). Different error sensitive areas in DEMs have been identified and visualised (Wood 1996, Maune 2001). Several approaches to explain and validate the errors have been conducted, standard procedures for checking data have been developed and some organisations have published their tests to accept or refuse DEM data from manufacturers (FGDC 1998, ISO-19113 2002). National geodata centres offer nowadays statistically reasonable information about the height accuracy of DEM.

All of these standards focus on a global view of the DEM. There is a gap between these indications and the use of the DEM in its derivations like slope, aspect or others. The error of adjoining cells or points is rarely investigated. Depending on the manufacturing process, this error is mostly indicated as “less than the global average” and therefore negligible. Goodchild (1996) stated a gap in the knowledge of the influence of the spatial relation of errors in DEM. He argued that the estimation of slope is directly dependent on the error of neighbouring points. The variance of the accuracy of slope would be given by

$$\sigma^2_{slope} = 2\sigma_e^2 \frac{(1-r)}{h^2} \qquad (1)$$

Where σ_e^2 is the variance of the height,

h = cell size

r = spatial autocorrelation (Goodchild 1986) of the neighbouring height errors

In practice, the correlation (r) would be nearby 1 (Goodchild 1996).

We investigated the correlation of errors of adjacent cells of DEM and its influence on derivatives in two test areas in and nearby the Swiss National Park. Different approaches were applied to describe error and uncertainty of height and slope to enhance the user’s knowledge about this source of uncertainty for wildlife studies.

2. METHODS

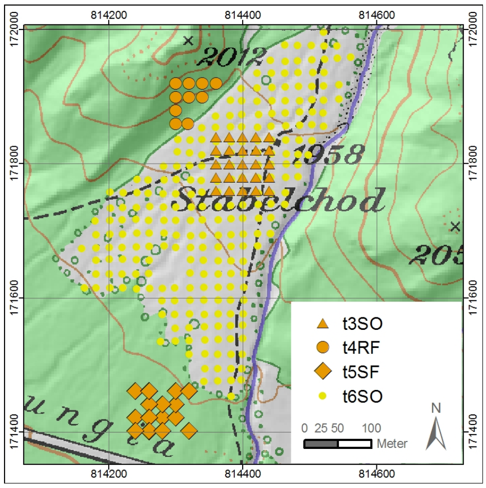
The two test areas were located in a mountainous area in the South East of Switzerland. 5 test sites were set up to distinguish different topographic characteristics. A rocky steep area and terraced grassland were chosen nearby Zernez in the Engadine valley (t1, t2), represented by a DEM with a resolution of 25 m provided by the Federal Office of Topography (Swisstopo 2004). It was originally produced semimanually from contour lines. In the Swiss National Park (SNP), a DEM with 20 m resolution was available. In contrary to the swisstopo DEM, this DEM was manually derived using an analogue photogrammetry station. Additionally to the height points every 20 m, breaklines were included. The second test area in the SNP (Stabelchod) was divided into 3 sites (t3, t4, t5). One of them was an open meadow with a gentle slope (SO), the second represents a relatively flat area in a forest (SF) and the third one was a fairly steep forested area (RF).

Figure 1. View of the test area „Stabelchod“ in the Swiss National Park.

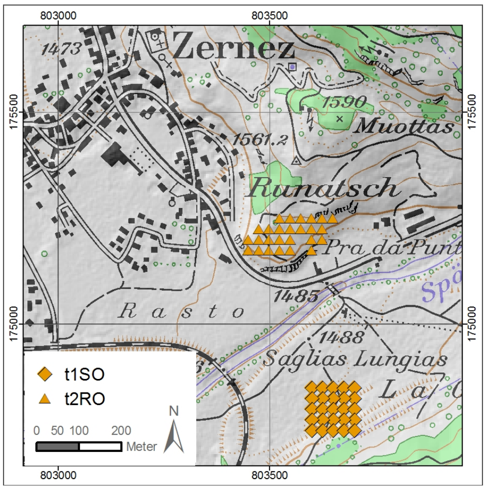
A forth set of control points includes the whole grassland of Stabelchod (t6SO, Figure 1). All in all, two test sites represented relatively rough terrain with high curvature and roughness, the other three had (at least for Swiss terms) relatively smooth terrain.

In 2000 the lattice points of the corresponding DEM were defined within a 5*5 test site and surveyed. The exact location of points in the field was determined by the authors using a surveying instrument (Leica TCA1102). The coordinates of the reference points were determined in 1998 with a surveying GPS (Trimble 4000 and 5000) with a horizontal accuracy of 0.7cm. The positional accuracy of the measurements of the surveyed points is within a few centimetres. Additionally a set of slightly irregular points over the Stabelchod area was used to enhance the number of available measurement points. This data defined exact vegetation plots for a wildlife study (Achermann 2000, Imfeld et al. 2006).

The location of 350 points, 53 plus 246 in Stabelchod and 51 in Zernez could be identified and surveyed. Some of the planned measurement points could be surveyed due to the inaccessibility of the area or the impaired visibility due to dense forest.



(a)



(b)

Figure 2. Overview and arrangement of the measurement points on the 2 test areas in the Swiss National Park (a) and the Engadine valley (b). (PK25 reproduction permitted by swisstopo)

The geometrical analysis including the geostatistics was done with ArcGIS 9.0 from ESRI and its programming environment.

A set of Macros (AML) was built for assessing the height accuracy of the lattice points and the DEM. The influence of the error on slope was done using GRID functions to define the neighbour cell with the highest height error.

3. RESULTS

3.1 Height error

Table 1 shows the summary of the height offset between reported (DEM) elevation and true elevation (survey). We calculated the absolute differences of height measurements. Only a few points in Stabelchod DEMs had height values below the true height, i.e. were underestimated by the DEM (Table 1 N<0). The normal case was an overestimation of the elevation compared with the true location height. The flat open areas of Stabelchod (t3SO and t6SO) show the lowest mean error with 0.59 m for t3SO 0.57 m and for t6SO. A slightly smaller error of 0.40 m was found in t5SF, a relatively flat forested area.. This last value, combined with a small RMSE (0.50) was realized in a forested area. A possible explanation for this is the DEM production which is based on manual photogrammetry and the habitat type of the forest. Mountain pine (*Pinus mugo*) dominates the area and the forest's degree of cover is only 60 – 90%. The highest average error was found in the steep forest area t4RF with 0.75 m and a RMSE of 0.95.

Table 1. Overall height accuracy of the 6 test sites and number of measurement points in the Engadine Valley and the Swiss National Park (S=smooth terrain, R=rough terrain, O=open meadow, F=forest)

	N	N<0	min	max	μ	σ	RMSE
t1SO	25	14	0.00	1.97	0.77	0.96	0.94
t2RO	26	3	0.00	10.72	2.54	2.66	3.45
t3SO	25	2	0.00	1.72	0.59	0.64	0.73
t4RF	10	2	0.16	1.82	0.75	0.74	0.95
t5SF	18	3	0.07	1.07	0.40	0.34	0.50
t6SO	246	51	0.00	2.99	0.57	0.60	0.75

The general average error of the photogrammetrically derived DEM is only slightly lower compared to the one based on contour lines, if the comparison is performed at the exact lattice point location (t6SO). Only the maximum error of 2.99 m in the t6SOset indicates that some small holes or bumps smaller than the resolution of the DEM were not correctly acquired in the 20 m resolution DEM.

The average error of the terraced grassland nearby Zernez (t1SO) was 0.77m and therefore slightly higher than the values of the sites in Stabelchod. but still much smaller compared to the rocky steep area t2RO with a mean error is 2.54 (n = 26, σ = 2.66, Max = 10.72 and RMSE = 3.45).

The spatial correlations of the errors are visualized in Figure 3. The small plots were showing a spatial correlation of the height error on local level. A different pattern is found in site T6SO with 246 points (Figure 3(f)). In addition to the local spatial error correlation, there seems to be a larger trend in a more global context, even though this can not be depicted from the corresponding semivariogram.

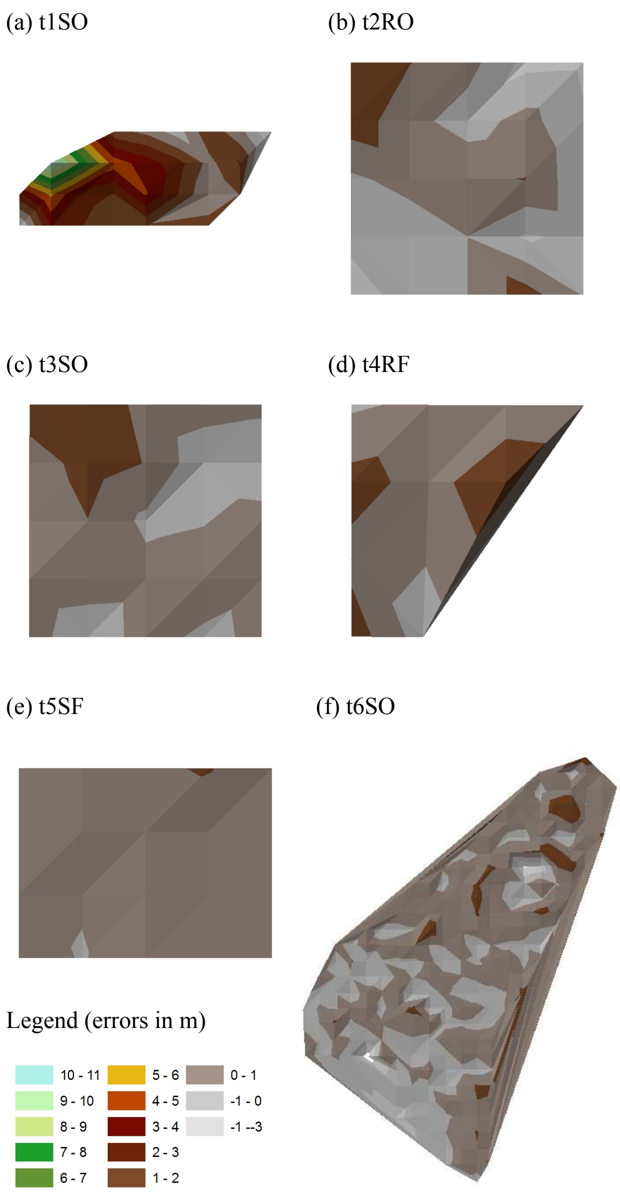


Figure 3. TIN with the measured errors at Zernez (a,b) and Stabelchod (c-f).

3.2 Slope error

To assess the influence of the height error on the slope derivative, the maximum difference of the height error in a 3*3 window around the centre cell was extracted. Here, the direction of the true value compared to the reported value is of interest. For example if a centre lattice point has a height error of +1m and the neighbour lattice point one of -1m, the difference between these to cells is 2 m. Figure shows the grid and the maximum error around each cell. Obviously there is a high spatial correlation of this maximum error. On t6SO the mean error between neighbouring cells amounts to 1.0 m (n = 245, σ = 0.5 m, min = 0.2 m, max = 2.91 m). On the rocky steep area the mean error was 4.3 m (n = 25, σ = 2.0m, min = 1.05 m, max = 8.14m).

Table 2 lists resultant errors in slope for the main figures above. The mean offset for the slope in the Stabelchod area based on t6SO is 2.89° while in the rocky area with breaklines in Zernez it amounts to 9.9°.

Table 2: Examples of errors of height with the corresponding error of slope (degrees) between neighbouring lattice points. The height errors represent empirically gained examples.

Height error of neighbouring cells	Cell size	Slope in °
0.22 m	20	0.63°
1.01 m	20	2.89°
1.91 m	20	5.48°
1.05 m	25	2.41°
4.3 m	25	9.90°
8.14 m	25	19.00°

The value for the autocorrelation for t6SO calculated with ArcInfo was 0.34 respectively 0.26 for t2RO.

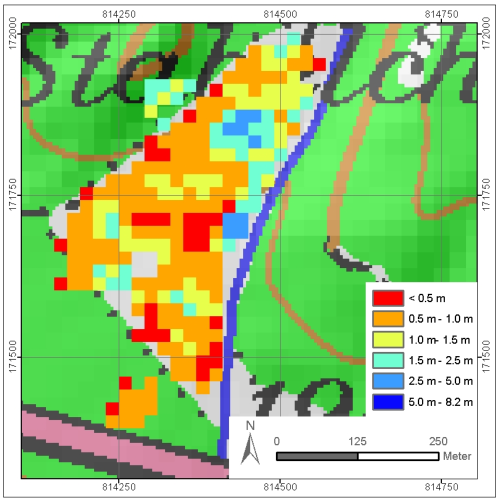


Figure 4. Distribution of the maximum error of neighbouring cells on Stabelchod, combining t4RF, t5SF and t6SO. (PK100 reproduction permitted by swisstopo)

4. DISCUSSION

The analysis has shown that the average error of the DEM in mountainous areas is acceptable for height values themselves, which are even lower than officially published. In its strict sense of altitudinal information, it is not of significance for wildlife studies. Swisstopo (2004) described a global average error of 4.1 m for the area of Zernez. We calculated 2.54 m and 0.77 m for the 2 corresponding test sites. This difference could be explained by the chosen control points of Swisstopo. They used survey points which are normally on ridges and exposed places. A constant underestimation in the DEM of these points often has been reported (Wood 1996, Martinoni 2001). Therefore, the global height error is even overestimated by the data provider.
The acceptance of global indicators of DEM accuracy as the sole measurements of data quality is dangerous. We showed that in our small test sites that the error is spatially dependent, but the autocorrelation is much lower than expected. The globally indicated height error provided by the data creator agency can be found even between 2 neighbouring lattice points. Therefore, the assumption that the local spatial autocorrelation is so high that these errors can be neglected for the calculation of derivatives like slope is not true for our study areas in this mountain region. In contrary, the potential error in the derived slope is surprisingly large that it needs to be considered when used in further modelling steps such as species habitat models. One method to deal with smaller amounts of inaccuracy in data is to reclass continuous variables into distinct classes. According to our results, such a classification has to be performed in a way that the influence of height errors becomes neglectible, i.e. the larger the error, the larger the class boundaries need to be.
Species distribution models often include aspect as one important factor as indicator of insolation, snow persistence or similar biologically important habitat factors. Especially in less steep areas errors in height can have large effects on the calculation of aspect, hence areas with slopes below a certain threshold, best derived from the local height error, should not be used for aspect calculations or marked as unreliable.
It might be argued that new technologies deriving global and local DEM are more accurate and therefore this kind of analysis might be obsolete. This increasing overall accuracy might be true, but in most cases this new technologies do increase the resolution of the cell as well. And users like wildlife ecologist will use this enhanced data for their studies. The uncertainty in slope and aspect calculations will therefore remain the same, if the resolution increases in the same extent as the accuracy of each single measurement, thus maintaining the problematic relationship for the calculation of derivatives. This means that there should be a clear focus to decrease the noise to signal ratio in DEM creation in strict relation to the resolution of the data, better horizontal resolution requires overproportionally highter vertial accuracy.

5. CONCLUSION

Often classes are built to distinguish the quality of the habitat. These classes should respect the height uncertainty and the low spatial correlation of errors.
Users should be aware that DEM quality questions not only base on manufacturing and global aspects. The planned application is very important and the adequate test for use is in the users responsibility.
More general, a shift from the providers data quality description to the user's test of fitness for use is indispensable. A local use of a global model like a DEM is not forbidden but on the users responsibility. Moreover, no data provider will deliver data quality descriptions for every possible or thinkable application. The costs are too high and the applications to numerous. The present study gave an example how users quite easily can improve their knowledge on height uncertainty.

REFERENCES

Achermann, G., 2000. The influence of red deer (*Cervus elaphus L.*) upon a subalpine grassland ecosystem in the Swiss National Park. Swiss Federal Institute of Technology, Zurich.

Aguilar, F. J., M. A. Aguilar and F. Agüera. 2007. Accuracy assessment of digital elevation models using a non-parametric approach. *International Journal of Geographical Information Science,* 21, pp. 667-686.

Campell, S., 2003. Winterökologie der Gemse (Rupicapra rup. rupicapra). Raumnutzung und Habitatwahl in einem Gebiet im Schweizerischen Nationalpark. Universität Basel, Basel.

Corsi, F., E. Duprè and L. Boitani, 1999. A Large-Scale Model of Wolf Distribution in Italy for Conservation Planning. *Conservation Biology,* 13, pp. 150-159.

FGDC, 1998. Geospatial positioning accuracy standards part 3: National standard for spatial data accuracy. FGDC-STD-007.3-1998, Federal Geographic Data Committee, Virginia.

Filli, F., R. Haller, M. Moritzi, M. Negri, J.-M. Obrecht, K. Robin and A. Schuster, 2000. Die Singvögel im Schweizerischen Nationalpark: Verbreitung anhand GIS-gestützter Habitatmodelle. *Jahresbericht der Naturforschenden Gesellschaft Graubünden*. Naturforschende Gesellschaft Graubünden, Chur, pp, 47-90

Goodchild, M. F., 1986. Spatial Autocorrelation, Concepts and Techniques in a modern Geography. *Geo Books*, Norwich.

Goodchild, M. F., 1996. The spatial data infrastructure of environmental modelling, pp. 11-15 *in* M. F. Goodchild, L. T. Steyaert, and P. O. Bradley, (eds.) *GIS and environmental modelling; progress and research issues*. Parker Dennison H., Fort Collins.

Guisan, A. and N. E. Zimmermann., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling,* 135, pp. 147-186.

Hirzel, A. H., J. Hausser, D. Chessel and N. Perrin. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data. *Ecology,* 83, pp. 2027-2036.

Imfeld, S., R. Haller and P. Laube, 2006. Positional Accuracy of Biological Research Data in GIS - A case Study in the Swiss National Park, pp. 275-280 in M. Caetano and M. Painho, (eds.) *7th International Symposium of Spatial Accuracy Asessment in Natural Resources and Environmental Sciences. Instituto Geográfico Português,* Lisbon, Portugal.

ISO-19113, 2002. 19113 - Geographic information - Quality principles (FDIS). Majorstuen.

Li, X., D. Li, Y. Li, Z. Ma and T. Zhai, 2002. Habitat evaluation for crested ibis: A GIS-based approach. *Ecological Research,* 17, pp. 565-573.

Martinoni, D., 2001. Models and experiments for quality handling in Digital Terrain Modelling. PhD. Geographical Institut University Zürich, Zürich.

Maune, D. F. (ed.), 2001. Digital elevation model technologies and application: The DEM users Manual. *American Society for Photogrammetry and Remote Sensing*, Bethesda.

McLoughlin, P. D., C. L. Ray, R. J. Gau, D. H. Cluff, R. Mulders and F. Messier, 2002. Hierarchical habitat selection by barren ground grizzly bears in the central Canadian Arctic. *Oecolgia*, 132, pp. 102-108.

Mladenoff, D. J. and T. A. Sickley, 1997. Assessing Potential gray Wolf Restoration in the northeastern United States: A spatial Prediction of favorable Habitat and potential Population Levels. *Journal of Wildlife Management*, 62, pp. 1-10.

Selkirk, S. W., 2002. Improving and extending home range and habitat analysis by integratino with a Geographic Information System. *Transactions in GIS,* 6, pp. 151-159.

Shi, W., P. Fisher and M. F. Goodchild, (eds.), 2002. *Spatial data quality*. Tylor & Francis, New York.

Swisstopo. 2004, DHM25 - Das digitale Höhenmodell der Schweiz. swisstopo, Wabern.

Wood, J. D, 1996. The geomorphical characterisation of Digital Elevation Models. PhD University of Leicester, Leicester.